Prosodic and multimodal markers of humor in conversation*

Salvatore Attardo, Lucy Pickering, and Amanda Baker
Texas A&M University / Georgia State University

This case study extends the findings of Pickering et al. 2009 to the domain of conversational humor. We find that, as was the case in humorous narratives, conversational humor is not marked by higher pitch or volume, increased speech rate, or significant pauses. Unlike narrative humor, conversational humor is not produced at a lower pitch and slower rate than non-humorous parts of the text. We find that smiling and laughter tend to occur with humor.

1. Introduction

In previous publications, Pickering et al. (2009) and Attardo and Pickering (2011), we have examined the suprasegmental prosody of joke telling. The results of the analyses of these short humorous narratives were unexpected, as they were found to contradict the "folk theory" of joke telling (for more details on the folk theory of humor performance and references, see Attardo and Pickering 2011). In a nutshell, the folk expectation was that the punch line of a joke, being the most significant part of a narrative, would be "marked" prosodically. In contrast, however, we found that, since punch lines occur at the end of a narrative, the placement of the punch line at the end of a prosodic unit "trumped" the functional markedness; thus, punch lines were not marked with unusually high pitch or volume. Other markers also did not allow the establishment of reliable correspondences between markers and humor (see below for more details).

While we consider these results very interesting, as they are among the first empirical studies of the prosody of humor, they are limited to narrative humor. In this paper, we expand the scope of our analyses to encompass discursive humor (broadly defined to include irony). Speakers can produce two types of humorous

turns in conversation: they can deliver a canned joke (a short narrative ending in a punch line[1]) usually by monopolizing the floor for the duration of the narrative,[2] or they can produce a conversational witticism. The difference between the two forms is distinct: a joke is a narrative largely, but not entirely, separate from the rest of the conversation, which contains a punch line, usually at the end of the text. The punch line builds on and exploits the narrative (usually called the "set up" of the joke). Conversely, the conversational witticism is not part of a separate narrative, but rather a comical "one-liner". Therefore the humorous line of the conversational witticism cannot build on a narrative and must either build on a previous turn by another speaker, or be interjected without prior preparation.[3] Typically, narrative canned jokes are introduced by negotiating sequences (such as "do you know this joke?") whereas conversational humor is typically not. To avoid confusion, we refer to the humorous part of a joke text as "punch line" and to a conversational witticism as "jab line" (more on this term below). We also avoid using "joke" for a non-narrative humorous text, although this is a common usage. On the differences between canned and conversational humor, see Attardo (1994: Ch. 10). Irony tends to be conversational humor, as it is rarely canned.

Table 1 summarizes the differences between jokes and conversational humor:

Table 1. Differences between jokes and conversational humor

Canned joke	Conversational humor
Narrative	Non-narrative
Punch line: humor at the end of a narrative	Jab line: humor anywhere in the text
Extended turn	Normal length turn
Introduced by negotiating sequence	Not introduced

The research question addressed in this study then is: are these differences between conversational and canned humor reflected prosodically? To put it differently, is conversational humor prosodically different from narrative, canned humor? Furthermore, it becomes interesting to consider how speakers interact around humorous turns, whether or not they choose to signal prosodically or kinesically the humorous nature of the turn and especially how they react to it (supportively or not).

Before we proceed with the discussion, we need to introduce two concepts from humor research: the jab line, mentioned above, and humor support.

1.1 The *jab line*

Just like the punch line indicates in humor theory the occurrence of a humorous instance at the end of the text (see Attardo et al. 1994 for evidence), the jab line

indicates the occurrence of a humorous instance anywhere else. Jab and punch lines are semantically indistinguishable, but they differ at a narratological level. Whereas punch lines are disruptive of the narrative they close, jab lines are not, and in fact often contribute to the development of the text (Attardo 2001). Punch lines and jab lines can be defined "broadly" or "narrowly". The broad definition is the sentence, turn, or pause-based unit in which the humor appears. The narrow definition is the (syntactic) phrase[4] within the sentence or turn in which the humor appears which removes completely the humor if deleted, all other things being equal. Broad and narrow definition may coincide, if the sentence or turn consists of only one phrase.[5]

1.2 Humor support

Humor support is defined as the conversational strategies used to acknowledge and support humorous utterances, among which figures prominently the production of more humor and/or laughter (Hay 2001). The strategies listed by Hay are:

- laughter (ibid.: 57–60)
- contribute more humor (mode adoption; ibid.: 60–62)
- echo (repeat part of the previous turn; ibid.: 63)
- offer sympathy or contradict self-deprecating humor (ibid.: 63–64)
- overlap and heightened involvement in conversation (ibid.: 65)

Hay further notes that humor support is not needed in at least two cases: for humor support itself and for irony (2001:65–66). Finally, Hay notes that the hearer may display understanding but not provide support, or withhold reaction entirely (and obviously these do not count as humor support).

It should be noted that Hay's original definition of humor support is limited only to vocal support, and so excludes smiles, for example. Hay was entirely aware of the fact that support can be also produced non-verbally, but chose to limit her analysis to the verbal text alone. There is obviously no need to retain that limitation and we will accept as humor support smiles and similar phenomena (e.g., smirks). We also add to the inventory the category of "metalinguistic humor support" to indicate all instances of comments on the humorous utterance that express support (most commonly asseverative particles such as "yeah!").

2. Methodology

2.1 Identifying humorous sequences

The analysis of conversational humor, or humor in discourse, faces a problem: the identification of the humor (specifically, telling apart the humorous parts of the text from those that are not). Most scholars use laughter and other known markers of humor. However, this approach suffers from a fundamental flaw: it systematically excludes from analysis those cases of humor that are not signaled by laughter (or other such markers). Moreover, speakers laugh even when nothing humorous is happening, for a variety of reasons (see Attardo 1994 for details). The standard technique employed to identify jokes, the Hockett test (removing phrases starting from the end of text and moving backward, until the humor disappears; Hockett 1973) does not work without modification in this context, because humor can occur anywhere in the text. Fortunately, the applications of semantically-based humor theory to long texts, such as short stories or novels, has encountered the same problem before and has handled it by simply scanning the text for script oppositions (i.e., parts of the text that are compatible with two opposed interpretations) and then performing a full scale semantic analysis of the fragment (see Attardo 2001). To put it differently, the text is scanned to see if any parts of it fit the criteria set out by humor theory as being humorous. In the case of discursive data we can utilize both semantic analysis and discursive markers such as laughter, smiling, etc. So, in a sense, our methodology to identify humor is a commonsensical triangulation using all the resources at hand, including humor theory.

2.2 Acoustical analysis

A significant problem in doing acoustic analyses of naturally occurring conversations is the quality of the data. The instrumentation (see below for the technical details) requires very "clean" recordings and any background noise degrades the capacity of the software to perform the pitch analyses on which scholars base their analyses. Recording high quality sound input often requires interfering with the setup of the situation (for example by requiring the speakers to wear lavaliere microphones or to speak into a microphone). While this may be obviated by using the recently introduced PZM (pressure zone) microphones, if one wants to have both audio and video recording, the setup is prohibitively expensive and cumbersome (multiple video cameras, synchronization of audio and video, etc.). Thus, we chose a different solution to these problems.

We recorded a dyadic exchange in which each participant was seated in front of a computer with a built-in camera and was wearing a headset/microphone

combination, of the type used commonly for chatting in online video-chats. This obviated the problem of audio and video recording, as well as providing a non-intrusive means of recording the exchange, since the subjects were all reasonably familiar with video chatting. Recordings were transferred to a Kay Pentax Computerized Speech Laboratory (CSL) Model 4300b in order to make instrumental measures. Pitch analysis was performed using the pitch contour analysis function of the CSL. Volume was determined by performing an energy analysis using the energy contour analysis function of the CSL. Pause length was calculated by selecting the area of the sound wave surrounding the pause, performing a spectrographic analysis and measuring the length of the pause manually using cursors. Speech rate was calculated by dividing the overall time taken to produce the pause-based unit, measured manually using cursors, by the number of syllables in the sample. Visual inspection was used to identify smiling in the videos.

The conversation was collected at a small Midwestern university in the US. Both the participants in the conversation were students who volunteered for the data collection and signed an IRB release. All the participants knew one another slightly from being enrolled in the same class. In this paper, we investigate a conversation between two young women, Carmen and Marina (these are pseudonyms). The conversation lasted slightly over five minutes.

The participants were given a set of instructions: they were given the text of the engineer and the frog joke, analyzed in Pickering et al. (2009), or another joke, selected because it does not end with a dialogic punch line.[6] The participants were told that they should tell one another their respective joke and that they should continue talking for five or ten minutes, until told to stop by the person in charge of the experiment. They were given the option of continuing to tell jokes or to talk about anything they wanted. No further instructions were given.

The prosodic transcription conventions are as follows: Following Brazil's (1997) model of discourse intonation and in the tradition of functionally based descriptions of English intonation the conversations are divided in tone units. Tone units boundaries are identified by the feature of prominence, i.e., fundamental frequency (F0) peaks which distinguish prominent syllables from the surrounding content. These prominent syllables are marked in CAPS (thus in line 1, the word OK, has the second syllable capitalized, to indicate its prominence) and are considered to be the linguistically significant features of the tone unit are measured for pitch and volume. Pitch and volume measurements for each prominent syllable are provided, in that order. Figures are in Hertz and decibel, respectively, and are in square brackets. A question mark indicates unreadable data. A range, such as [232–204] on line 292, indicates a falling tone. Double forward slashes indicate pauses and these pause-based units may comprise one or more tone unit. Pauses are measured in seconds. All pauses are transcribed on a separate line, for ease

of identification. So, for example, line two has the first pause of the conversation, which lasts 0.15 seconds (i.e., 15 hundredths of a second).

1	C	//UM [204][63] oKEY [225][69]//
2		0.15
3	C	// THERE [139][68] we GO [204][63] //

If a pause is followed by a notation in double round brackets, such as ((smack)), this indicates that somewhere during the pause, there occurred a lip smacking sound. Thus on line 166, we see a 2.5 seconds pause which includes a lip smacking, within a turn by Carmen:

165	C	//YEAH [270][68] //
166		2.5 ((lip smack))
167	C	//YUP // [230] [60]

The numbering of the lines is for convenience of reference and has no theoretical value. Author comments and impressionistic notes are in curly braces. Overlapping turns are marked with a "=" in each speaker's turn. The overlaps are not typographically marked by indentation. Other conversation analytical typographical conventions are followed, such as marking lengthened vowels with colons.

2.3 A note on the coding of humor and humor support

Generally, the production of humor to support humor is a well-documented phenomenon. However, when coding a turn for analysis it is important to distinguish between humorous turns of humor supports and non-humorous turns of humor support. Consider the following exchange:

286	M	// you KNOW [147][65] = //
287	C	// =you're aMERican [227][60] //
288		1.03
289		((both laugh))
290	M	// no KIdding [245][75] //
291		0.42
292	C	// YEAH [232–204] [71] (M. laughing) //
293		1.48
294	M	// DAMN [256][58] Americans //
295		0.09
296	C	//YESS [no Hz][43–71] //

On line 287, Carmen interrupts Marina with a jab, recycling the punch line of the old joke about Americans being monolingual.[7] The joke is followed by a pause and by shared laughter, which is a form of support (Hay 2001). Then Marina produces an ironical supportive turn: "No kidding". This is followed by a supportive turn by Carmen "Yeah". This turn is not humorous, while being supportive. Marina laughs, expressing once more support and after a pause produces another jab line "Damn Americans!" which is once more ironical. Carmen provides non-humorous support, on line 296. So, in this exchange, Carmen supports Marina's humor non-humorously, by providing asseverative particles. Marina, conversely, supports Carmen's humorous turn in 287 by laughing in 289 and in 292, and by producing two ironical turns (290; 294). So, in conclusion, it is important to analyze carefully humor support, since not all humor support is humorous.

3. Overall results

We found thirteen instances of humor in the conversation.[8] Two were the punch lines of the canned jokes the speakers were asked to perform. The remaining eleven instances of humor break down as follows: five jab lines (conversational witticisms), three cases of irony, and three cases of irony support.

3.1 Speech rate

A comparison of the speakers' speech rate (turn duration divided by number of syllables in the pause-based unit) shows that humor is not marked by a different speech rate: even a superficial observation of Table 1 shows that whereas Carmen

Table 1. Speech Rate (rounded to two decimals)

Speaker		Humor	Serious	Total
Carmen	Average of Speech Rate	0.19	0.31	0.25
	Average Number of Syllables	7.83	5.02	6.425
	Average Turn Duration (in seconds)	1.47	1.06	1.265
Marina	Average of Speech Rate	0.27	0.23	0.25
	Average Number of Syllables	4.14	4.56	4.35
	Average Turn Duration (in seconds)	1.09	0.86	0.975
Total Average of Speech rate		0.23	0.27	0.27
Total Average Number of Syllables		5.84	4.80	4.88
Total Average Turn Duration (in seconds)		1.27	0.97	0.99

Table 2. Speech Rate, across genres of humor

	Irony	Jab	Narrative	Serious
Carmen	0.23	0.16	0.17	0.31
Marina	0.30	0.26	0.16	0.23
Average	0.265	0.21	0.165	0.27

has a slower rate in humor, Marina has a faster one. Statistically these differences are not significant (Paired T test, $p = 0.7$).

We checked whether there was a difference once we separated irony, narrative punch lines, and jab lines, contrasted with serious speech rate (Table 2), but once more the results were not significant (Repeated Measures Analysis of Variance, $p = 0.37$).

3.2 Pauses

We find that the mean length of pauses in the data was 0.70 seconds. Only one of the four instances of humor that occurred with a within-turn-pause has a length that was above the average pause and that by a mere .07 seconds. Three instances of humor occurred as overlapping turns, and hence had no pause at all preceding them. Hence it follows that at least half of the instances of humor were not preceded by significant pauses. However, it would be misleading to think that the remaining instances of humor were examples that support the theory that an instance of humor is preceded by a pause that marks it, since they were all instances of pauses between turns. Two considerations arise: 1) pauses between turns are significantly longer that within-turn pauses, and 2) pauses between turns cannot be interpreted as "marking" the following turn, unless the next speaker has been selected in the previous turn. Finally, we should point out that none of the pauses occur before the punch or jab line in the strict definition. In the broad definition (i.e., the pause-based unit that contains the punch or jab line), all pause-based units containing instances of humor will be preceded by a pause; therefore, it would be meaningless to argue that our data support the pause-marking hypothesis. Only pauses that precede immediately the phrase containing the punch/jab line, above average length within-turn pauses, or exceptionally long pauses between turns could count as evidence of marking of the punch or jab line. Under these restrictions, we found that none of the instances of humor was preceded by a signaling/marking pause. It should be noted that Wennerstrom and Siegel (2003) found an average pause duration slightly under 0.75 seconds between syntactically complete units. Furthermore, Wennerstrom and Siegel (2003) report that pauses

of 1.5 seconds are not unusual between turns. Hence, the pauses in our conversation are not exceptional.

3.3 Pitch

We found that our two speakers tended to produce humorous turns at a very slightly higher pitch than serious turns, on average. The difference between humorous and non-humorous turns was not significant. The differences between serious turns and punch lines, jab lines, and irony taken individually were likewise not significant. The differences between speakers were also insignificant. See Table 3/4, Hertz values. The zero figures in Carmen's support reflect the fact that she did not produce any.

Table 3. Comparison of humorous vs. serious turns, pitch and volume

		Carmen	Marina	Both speakers
Humor	Average of Hz	214	205.28	209.64
	Average of Db	64.33	70.42	67.37
Serious	Average of Hz	216.49	204.60	210.54
	Average of Db	66.94	69.78	68.36
Total Average of Hz		215.24	204.94	210.09
Total Average of Db		65.63	70.1	67.86

Table 4. Pitch and Volume results, across genres of humor.

		Speaker		
Type of humor		Carmen	Marina	Both speakers
irony	Average of Hz	227	175	201
	Average of Db	62	68	65
jab	Average of Hz	213.25	199.66	206.455
	Average of Db	64.25	73.33	68.79
punch line	Average of Hz	204	162	183
	Average of Db	67	72	69.5
support	Average of Hz	0	250.5	250.5
	Average of Db	0	66.5	66.5
Total Average of Hz		214.75	196.79	205.77
Total Average of Db		64.42	69.95	67.19

3.4 Volume

The present data (Table 3/4, decibel data) show that punch lines were produced at a slightly higher volume: we find that the humorous turns are produced on average 0.99 dB higher than serious turns. The difference is insignificant. We did however find a difference between the speakers' serious turns, which differ on average by 3 dB; to put it differently, Marina speaks a little louder than Carmen, but only when they are not being funny. No significant differences were found among types of humor, even after separating the two speakers' data.

3.5 The canned jokes' performance

The next step in our analysis is to check whether Carmen and Marina differ from the joke tellers in Pickering et al. (2009). We can briefly summarize the results from Pickering et al. (2009) that are relevant to the present study. Using a sample of ten speakers each performing two jokes, one provided to the participants and one improvised by them, we found that

- Punch lines were significantly lower in pitch than the set-up of the joke (we confirmed this finding using semitone data rather than raw F0 measurements; on average punch lines were 1.73 semitones lower than the setup).
- Punch lines were produced at slightly higher amplitude which was on average 0.2 dB higher in volume; however the difference was not statistically significant.
- Punch lines were not preceded by significant pauses; in fact, 15 jokes (75% of the texts) did not have a significant pause before the punch line.
- Speakers tended to deliver the punch line of the joke at a significantly slower rate than the set-up in both types of jokes.
- 60% of jokes had either laughter or smiling voice or both, in the punch line.

4. Comparison of the results for the dyadic exchange vs. the canned performance data

4.1 Pitch

From even a cursory examination of Tables 3 and 4, it is obvious that the results of the conversational data analyzed in this study are consistent with those collected in the interview setting of Pickering et al. (2009). We find that the average pitch of the non-humorous prominent syllables do not differ significantly from the average humorous ones. Furthermore, neither the punch lines taken by themselves,

234 Salvatore Attardo, Lucy Pickering, and Amanda Baker

nor the jab lines, or the ironical turns, differ significantly from the average pitch. There are some insignificant but suggestive data, such as the fact that ironical turns tend to be slightly higher than the average pitch, but no general conclusion can be drawn on this basis.

The most obvious difference with the results in Pickering et al. (2009) is that in these examples, we no longer find significantly lower pitch for the humorous sequences. In Pickering et al. 2009, we propose that the significantly lower pitch found in punch lines of the humorous narrative reflects the importance of the larger prosodic structure of the paratone. Also known as speech paragraphs and pitch sequences, paratones are a prosodic unit comprising tone units that begins with a high pitch and ends with a final low pitch, which may also be accompanied by longer pauses. Paratones have been identified across genres within longer sections of spoken discourse, and function in a way similar to paragraphs in written discourse. In the case of our data, the final closing low pitch of the paratone "trumped" the possible higher pitch that might be expected for the punch line. In the data discussed here, however, only two of the thirteen instances of humor are punch lines, whereas all the instances of humor in Pickering et al. (2009) are punch lines. Hence the results in this study are consistent with the explanation that the position at the end of a paratone explains the significantly lower pitch for the punch lines. Out of the two punch lines in this study, one is much lower (40Hz) than the serious text, while the other is in fact higher than the serious text. Given that we are considering only one instance, it is impossible to further speculate on the significance of this fact.

We then tested whether the humorous turns might not be average in pitch, in relation to the conversation at large, but nonetheless significantly marked in relation to the co-text in which they occur. In order to test this hypothesis, we compared the turn immediately before the humorous turn with the humorous turn itself. In some cases there were two prominent syllables in the humorous turn and we compared the second one, which was in both cases the one closest to the jab/punch line. In two cases we had to discard a humorous turn, which occurred immediately after another humorous turn (humor support). The results show that there is no significant difference in pitch (the humor is on average 8 Hz higher than the preceding turn). The actual values can be found in Table 5 below.

4.2 Volume

The present study is consistent with the result in Pickering et al. (2009): humorous turns are produced on average 0.99 dB higher than serious turns (0.2 dB in Pickering et al. 2009). This difference, as it was in Pickering et al. (2009), is insignificant. We conclude that canned jokes and conversational jokes do not differ in terms of volume marking of the humor. Like for pitch, we checked if humor may differ in

All rights reserved

Table 5. Comparison between humorous turns and their preceding turns.

Pitch (measured in Hz)		Volume (measured in dB)	
Preceding turn	Humor	Preceding turn	Humor
224	162	73	72
200	216	69	74
216	204	74	74
232	204	64	67
172	179	68	72
218	227	59	62
147	227	65	60
218	256	71	58
162	185	67	63
Total average	Total average	Total average	Total average
199	207	68	67

volume from the preceding turn, but we found no significant difference in volume, with the humor being one decibel softer than the preceding turn, on average (Table 5, Decibel measures).

4.3 Speech rate

As we saw, the present study found no significant difference in the speech rate of humorous and serious turns. This is at variance with the results of Pickering et al. (2009) where we have found that punch lines were delivered at a significantly slower rate than the surrounding text. We return to this datum below.

5. Comparison of conversational humor vs. canned humor

An issue of importance in the pragmatics of humor has been the differences between conversational/discursive humor and narrative/canned humor. In this context, we are interested in the prosodic differences between the two, if any.

5.1 Speech rate

We noted above that the results of the present study failed to duplicate the significantly lower speech rate of punch lines, found in Pickering et al. (2009). We found instead no significant difference between serious and humorous turns' speech rate.

However, we note that the narrative punch lines were delivered at the slowest speech rate of all the humorous events (0.165 on average; see Table 2). Conversely, irony and serious turns are essentially delivered at the same speech rate (0.265 and 0.27, respectively, see Table 2). Hence, we hypothesize that there may be a correlation between narrative punch lines and lower speech rate, absent in irony and jab lines. This may be then a difference between narrative and non-narrative humor delivery.

5.2 Volume

As we saw above (Table 5), we found no significant difference in volume, neither between serious and humorous turns, nor between types of humorous turns. Hence we conclude that narrative and discursive humor do not differ significantly in terms of volume.

5.3 Pitch

As we saw above (Table 5), we found no significant difference in terms of pitch between serious and humorous data. This is at variance with the significantly lower pitch found for punch lines in Pickering et al. (2009). This fact may be explained, as we hypothesize above, by the fact that punch lines occur by definition at the end of a paratone, and hence show a lower pitch. The present data do not allow us to generalize any further, so, while this may be a difference between narrative and discursive humor, we cannot make any conclusive claims. Further research is being conducted to answer this question.

5.4 Pauses

We will now consider pauses. One of the most interesting conclusions of our previous studies is that punch lines of canned jokes are not generally preceded by a significant pause. We find an example of a joke not preceded by a pause in our text:

276	C	// do you speak any OTher [225][62] languages besides ENGlish[242][60]? //
277		1.11
278	M	// NOPE [251][71] //
279		0.42
280	M	// nothin FLUently [216][59] //
281		0.12
282	M	// little bit a this n THAT [210][60] n //
283		1.26
284	M	// k OTher [185][73] thing //

285		0.66
286	M	// you KNOW [147][65] = //
287	C	// =you're aMERican [227][60] //
288		1.03
289		((both laugh))

Here Carmen interrupts Marina's turn to joke about the well-known stereotype that Americans are monolingual. Clearly, no pause precedes jab line since it is an interruption. We examine elsewhere the humor support of this passage. The jab line on 257 is also an interruption. We can thus add the notion of "interruptive jab line" to the toolbox of the prosodic analysis of humorous discourse, which by definition is not preceded by any pause at all.

Other jab lines seem more complex. Consider the following exchange:

140	C	//um what should we TALK [140?][65] about?//
141		1.042
142	M	//I have no iDEa [210][68]//
143		0.393
144	C	//oKAY [250][65]//
145		1.39
146	C	// there's lots of BOOKS [243][66] in here//
147		1.57
148	M	//yeah I'm LOOKin [250][68] around at all these PApers [188][64]//
149		0.34
150	M	//hangin on the WALLS [172][65] //
151		0.52
152	M	//this OFfice [198][71] is really DIRty [182][67]//
153		0.64 ((smack))
154	C	// this OFfice [294–279][67] is really CLEAN [222][62] //
155		0.44
156	C	//i THINK [334][63] it's Steve BROWN's [206][58]//
157		0.92
158	M	//AHH. [172][68]//
159		0.52
160	C	//yeah//
161		0.883
162	M	//this looks like SAL's [179][72] //
163		0.21
164	M	((laughs))

The passage occurs immediately after an exchange in which Marina and Carmen agree on the fact that they have nothing obvious to tell one another. Carmen volunteers an observation about her surroundings. Now it is necessary to explain that the recordings were done in one of the authors' office (Salvatore Attardo) and the office of one of his colleagues, Steve Brown, who taught and had an office in the same department. Both students knew both faculty members. Carmen was in Brown's office and Marina was in Attardo's office. So, Carmen is fishing for a topic, as can be deduced by the 1.39 second pause, and hits on the fact that there are many books in the office. Marina doesn't respond until 1.57 seconds later, clearly also fishing for something to talk about (as deduced from the two pauses of .34 and .52 seconds). So, Marina describes the papers hanging on the walls of the office and concludes "This office is really dirty". Carmen matches the syntax but not the prosody of the previous turn and says: "This office is really clean". So far, there is no humor at all in the conversation segment. Then after a .44 pause, Carmen continues stating her belief that the clean office belongs to Steve Brown. Consider now the humorous turn "This looks like Sal's" in which Marina creates a humorous comparison between the clean office and the dirty office and by implication, their owners. That Marina intends this turn as humorous is made explicit by her laughter at the end of the turn.

The jab line "This looks like Sal" is preceded by a pause (0.83). This would then seem to be an example of a jab line preceded by a pause. However the pause occurs between turns. Conversely, we find a within-turn-pause (.44) on line 155. It seems fairly obvious that the kind of pauses we are interested in are within-turn pauses (pauses in joke telling monologues are obviously within-turn pauses). So, when considering dialogical data, we must differentiate between within-turn pauses and inter-turn pauses. Moreover, we need to point out a fairly obvious point: since we are using pause-based units as our unit of analysis, it follows that *all our units are preceded by a pause*. So, if the notion of pauses somehow "marking" humor or punch lines has to retain any content, we need to specify that the pause should immediately precede the jab line, applying the Hockett test to the pause based unit. In the example above, the "narrow" definition of jab line clearly does not show a pause preceding the jab line. It is important to repeat this point, already made above: since the units of analysis are delimited by pauses, if we merely state the hypothesis as "a (broad definition) jab line is preceded by a pause", then the definition is tautologically true, since all pause based units are preceded by pauses. Hence, we either restrict the pauses under investigation to within-turn pauses, or we adopt a narrow definition of jab line, or both.[9]

Consider now the example of humor support after "you're (an) American". In this case the jab falls flat as it is followed by a long pause. However, Marina then follows up first with a turn expressing agreement and finally with a turn in which

she adopts the humorous mode, by uttering the ironical "Damn Americans" turn. This can be seen as expressing humor support, by partially repeating the jab line "You're American". Significantly, no within turn pause appears before the humorous turn "Damn Americans". However, since this is an example of irony, it is fair to note that we are not aware of any claims that irony should be preceded by a pause. There are two instances of irony preceded by a pause in the data, but it seems more likely that there are not "pregnant" pauses, but rather hesitation pauses, as suggested by the fact that they co-occur with filled pauses (um, hh).

In conclusion, we must restate what said above: pauses do not mark punch lines, jab lines, or irony in any significant way. What pauses are seen to occur near or after humor instances are best interpreted as hesitation or other dysfluencies. Of course, it is possible that some instances of humor may be found that are preceded by significant, emphatic pauses, but that does not mean that pauses can be taken to mark humor in general.

6.   Non-prosodic markers

We have so far deliberately limited our discussion to linguistic/prosodic features, with the exception of the mention of smiling as humor support above. However, in Pickering et al. (2009) we examined smiling and laughter as potential markers of the humor in the text. Our conclusion in that article was the smiling and laughter could not serve as reliable indicators, since they occurred in about 50% of the cases. We now turn to the examination of the present data, to see if there are any differences.

The reception of humor is particularly interesting: the jab discussed above (162) is followed immediately by the following turn:

162	M	//this looks like SAL's [179][72] //
163		0.21
164	M	((laughs))
165	C	//YEAH [270][68] //
166		2.5 ((lip smack)) h h
167	C	//YUP [230] [60]//

in which Carmen seemingly rejects Marina's attempt at framing the remark as humorous and follows up with a 2.5 second pause, followed by a variant repetition of "yes" and an even longer pause. The transcription of the conversation is here useless, and indeed even a careful spectrographic analysis of the passage reveals only two aspirations (marked by "h" in the transcript). It is only by examining the video re-

cording that we can ascertain that Carmen produces two small smiles that coincide with the aspirations, which can be then interpreted as reduced forms of laughter.

Figure 1. Carmen-smiling around line 166.

This observation brings up an interesting point to consider: those studies that utilize only the transcription of the conversational exchange would have recorded this as either failed humor (since Marina signals her intention to produce a humorous turn by laughing after it, line 164) or as no humor at all, since it is not followed by audible laughter. We can only conclude that only a multimodal analysis can be completely reliable in this kind of issue.

All instances of humor that occur in the conversation (n = 13), except one, are supported and/or acknowledged by laughter (see Table 6). The following is the lone unsupported, and hence presumably failed, instance of humor.

65	C	// here's MY [216][74]joke //
66		0.62
67	C	//um=//
68	M	// =do not use big WORDS [204][74] //
69		0.08
70	C	//i don't reMEMber [188][71] my joke //
71		2.48
72	C	//ok ONCE [242][74] upon a time there was a engiNEER? [194][62] //

On line 68, Marina tries to introduce a self-deprecating jab line, the obvious inference being that Carmen should not use big words because they are too difficult to understand. However, Carmen completely fails to acknowledge the attempt (she continues to produce the sentence "Um (…) I don't remember my joke"). Incidentally, since Marina interrupts the filled pause in line 68, there is obviously no pause prior to the jab line, in this case either.

Consider now the overall results, reproduced in Table 6. We immediately note the preponderance of smiling and laughter across the table: twenty-one instances

Table 6. Facial expression at and immediately following humor instances (shaded cells represent the occurrence of the humor).

Line number	Humor	Speaker	Type	Facial Expression	
				Marina	Carmen
46	Was a donkey	M	Punch	Lips pulled tight, zygomatic lift	Face invisible maybe smile
50	Hahaha	C	Irony	smile	Smile, laughter
56	Your response [face]	M	Jab	Smile, laughter	Smile
60	I apologize	C	Irony	smile	Smile
68	Do not use big words	M	Jab	smile	Smile
113	Talking frogs are really cool	C	Punch	smile	Neutral, followed by smile
162	This looks like Sal's	M	Jab	Laughter (face averted from screen)	Smile
214	Well, this is fun	C	Irony	Lips pulled tight	Light smile
216	hmm very	M	irony support	Smile (non-Duchenne)	Light smile
287	you're (an) American	C	Jab	Neutral (surprise?) followed by smile and laughter	Smile, cocks head
290	No kidding	M	irony support	laughter	Laughter
294	Damn Americans	M	irony support	smile	Smile, downward gaze, shakes head
355	Me and these headphones are not getting along	C	Jab	Neutral, followed by smile and laughter	Face invisible

out of a possible twenty-six. It is obvious that speaker and hearer collaborate in reassuring one another that the humorous turns are produced and interpreted as humorous. In a few cases, either because the speaker turned away from the camera, or because their gaze was directed downward, it was impossible to see what the speaker was doing (although in one case her laughter is nonetheless audible). If we discard those cases, we are left with an interesting observation: both cases of punch line are delivered without a (full) smile, but are received with a smile. All other instances of humor for which we have data are delivered with a smile or laughter. This could then be a difference between canned, narrative humor and conversational

Figure 2. Marina's zygomaticus contraction

humor: the former can (must?) be delivered with a deadpan expression and can only be followed by smiling/laughter. All other conversational humor (jab lines and especially irony) is delivered with at least a smile. It is particularly interesting that the only two non-Duchenne[10] smiles (contraction of the zygomatic muscle) occur in Marina's facial expressions: the first one occurs in the delivery of the punch line of her joke, which is delivered with a tight-lipped expression, with zygomatic contraction and probably caninus contraction as well (see Figure 2). The second one occurs in her reaction to the irony support in line 216.

This leads us to reiterate the consideration we expressed before concerning the usefulness of multimodal analysis: so far, we have identified only one fairly reliable marker of humorous intention, a smiling or laughing expression, with the proviso that potentially canned, narrative jokes are the exception to this.

6.1 Liminal sequences

A potential objection to the claim that our data represent authentic unscripted conversational data is that we did give the participants some direction. However, the directions concerned exclusively the telling of the two canned jokes, necessary to establish a baseline and compare the conversational data to the "performed" data in Pickering et al. (2009). In any case, it is very clear that Carmen and Marina are aware of the difference in status between the "required" part of the conversation and the unscripted part. After laughing at the second joke the following exchange takes place:

123		2.41
124	M	// oh well //
125		0.05
126	M	//so what's UP? [193/162?][71]//
127		0.408
128	C	//um //

129		0.28
130	C	//NOthing [200][62] //
131		2.51

The presence of pauses for a total of more than five seconds of silence clearly shows the liminal nature of this sequence. Marina tries to get the conversation part of the assignment going, by using the conversational formula "what's up?" but her opening is rejected by Carmen, since her question is followed by "Nothing" and a long pause. Interestingly, Marina does not take advantage of the long pause to take a turn. The "conversation" continues:

132	C	//YOU? [235][?]//
133		1.48
134	M	//no::t a THING. [325?][66]//
135		0.48 ((smack))
136	C	//oKAY [180][67] //
137		1.41
138	C	//SO [270][71] //
139		0.26
140	C	//um what should we TALK [140?][65] about?//
141		1.042

Marina waits almost two seconds to parallel Carmen's rejection ("Not a thing") signaling her refusal to be the one to start the conversation (we may speculate that this is so, perhaps for fear of volunteering personal information). Carmen seems to realize that Marina is unlikely to start the conversation and, since after all, Carmen was the first to reject Marina's earnest attempt to get it going, she closes the interaction "Okay, so" and moves on by "going meta" so to speak, i.e., initiating a discussion of what the discussion should be about. Here we witness key matching as an affiliative move. On line 136, Carmen uses a marked intonation "hump intonation" which Marina had used in her reaction to Carmen's joke. Marina picks up the same marked "hump intonation" in line 142. We interpret picking up and repeating a marked key as an affiliative pitch concord (Brazil 1985; Wichmann 2000: 141) or prosodic orientation (Szczepek Reed 2006: 34).

140	C	//um what should we TALK [140?][65] about?//
141		1.042
142	M	//I have no iDEa [210][68]//

Indeed, after this turn, Carmen and Marina start chatting and continue to do so until the data collector interrupts them.

7. Conclusions

In conclusion, we have found several interesting results: most significantly that the conversational humor was not produced with significant prosodic differences from serious text and that no significant differences were found between narrative, conversational, and ironical humor, as far as volume, pitch, speech rate, and pauses were concerned. However, when comparing the results for narrative/canned humor to the conversational data we found that, whereas narrative humor's punch lines were significantly lower in pitch and speech rate than the serious text (Pickering et al. 2009), this does not hold for conversational humor. This is seen as a confirmation of our prior hypothesis, expressed in Pickering et al. (2009), that the lower pitch of narrative humor's punch lines is due to the location of the punch lines at the end of a narrative and hence at the end of a paratone.

Overall then we can say that the hypothesis that conversational and narrative/canned humor differ significantly, as far as the prosodic markers we have examined, has been refuted and particularly so the idea that humor is signaled by pauses or emphatic prosody. This is particularly effectively represented by interruptive humor that shows contextually appropriate unremarkable prosody. Furthermore, we have shown that a multimodal analysis is crucial, in some cases, for the proper interpretation of the data and that facial expression, as well as laughter, is a significant marker of humor support. Needless to say, while this is the first study ever to compare canned and conversational humor prosodic patterns, we are very much aware of the fact that our findings are based on a small sample and therefore they must be interpreted as preliminary and tentative. Further research will broaden the size of the sample and therefore allow for checking of statistical reliability and address other variables, such as the potential difference between trained (professional comedians) and non-trained performers.

Transcription Chart

((lip smack))	Paralinguistic events
[204][63]	Pitch and volume values for a prominent syllable, in Hertz and Decibel, respectively
[?]	Impossible to get an accurate measurement

CAPS	Prominent syllable
// ... //	Beginning and end of pause-based unit
0.48	Pause duration in seconds
h h	aspiration
=	Overlapping turns
{ ... }	Beginning and end of author comments
:	Lengthened vowel

Acknowledgements

Notes

1. There are exceptions: some jokes are not short, some of them are not (entirely) narrative, and some do not end with a punch line (for example, shaggy dog jokes, or meta-jokes). We can safely ignore these details in this context. The interested reader should refer to Attardo (1994) for discussion and references.

2. This is not to say that the hearer(s) cannot interrupt or provide backchannel support.

3. Conceivably, the speaker could use his/her own prior turn as a setup of sorts and present a punch line at the end of his/her turn; we would categorize this as a narrative joke.

4. The use of phrases (NP, VP, PP, etc.) in the Hockett test (see below) to segment the sentences of jokes and locate punch lines was adopted in Attardo et al. (1994). The only difference in the present context is the extension of the method to spoken utterances, as opposed to written ones.

5. One of the referees notes that the broad and narrow definition of jab line correspond to syntactic and prosodic criteria defining the domain of the humorous instance.

6. We do not consider the results of the dialogic/non-dialogic punch lines in this article. They will be the subject of ulterior publications.

7. What do you call someone who speaks three languages? Trilingual. What do you call someone who speaks two languages? Bilingual. How about someone who speaks only one language? An American.

8. Given the relatively small sample and the fact that the data come from the same conversation, no claim is made that the results can be generalized or are statistically valid. Statistical calculations are given merely as preliminary analyses of a case study.

9. We should note another possible line of argument to differentiate between pauses: some researchers (e.g., Grosjean 1980) have suggested that the perceived length of pauses is affected by their syntactic placement, i.e., for example within constituent or between constituents. One could perhaps differentiate between within-turn and intra-turn pauses on the basis of relative saliency of the pause. We will not pursue this line of argument in this context.

10. On non-Duchenne smiles and the Facial Action Coding System, see Ekman and Rosenberg (1997), especially the chapters by Ruch.

References

Attardo, S. 1994. *Linguistic Theories of Humor*. Berlin: Mouton De Gruyter.

Attardo, S. 2001. *Humorous Texts*. Berlin: Mouton De Gruyter.

Attardo, S., Attardo, D.H., Baltes, P., and Petray, M.J. 1994. "The linear organization of jokes: Statistical analysis of two thousand texts". *HUMOR: International Journal of Humor Research* 7(1): 27–54.

Attardo, S. and Pickering, L. 2011. *"Timing in the performance of joke"*. *HUMOR: International Journal of Humor Research* 24(2): 233–250.

Brazil, D. 1997. *The communicative Value of Intonation in English*. London: Cambridge University Press.

Ekman, P. and Rosenberg, E.L. 1997. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford: Oxford University Press.

Grosjean, F. 1980. "Linguistic structures and performance structures: Studies in pause distribution". In H.W. Dechert and M. Raupach (eds), *Temporal Variables in Speech*. Berlin: Mouton De Gruyter, 91–106.

Gustafson-Capková, S. and Megyesi, B. 2001. "A comparative study of pauses in dialogues and read speech". In *Proceedings of Eurospeech 2001*, Volume 2, Aalborg, Denmark, September 3–7, 2001: 931–935.

Hay, J. 2001. "The pragmatics of humor support". *HUMOR: International Journal of Humor Research* 14(1): 55–82.

Hockett, C.F. 1973. "Jokes". In M.E. Smith (ed), *Studies in Linguistics in Honor of George L. Trager*. The Hague: Mouton, 153–178. Reprinted in his 1997, *The View from the Language*. Athens, GA: University of Georgia, 257–289.

Jefferson, G. 1989. "Preliminary notes on a prosodic metric which provides for a 'standard maximum' silence of approximately one second in conversation". In D. Roger and P. Bull (eds), *Conversation: An Interdisciplinary Perspective*. Clevedon: Multilingual Matters, 166–96.

Pickering, L., Corduas, M., Eisterhold, J., Seifried, B., Eggleston, A., and Attardo, S. 2009. "Prosodic markers of saliency in humorous narratives". *Discourse Processes* 46: 517–540.

Roberts, F., Francis, A.L., and Morgan, M. 2006. "The interaction of inter-turn silence with prosodic cues in listener perceptions of 'trouble' in conversation". *Speech Communication* 48 (9): 1079–1093.

Szczepek-Reed, B. 2006. *Prosodic Orientation in English Conversation*. New York: Palgrave Macmillan.

Thomason, W.R. and Hopper, R. 1992. "Pauses, transition relevance, and speaker change". *Human Communication Research* 18(3): 429–444.

Trimboli, C. and Walker, M.B. 1984. "Switching pauses in cooperative and competitive conversations". *Journal of Experimental Social Psychology* 20 (4): 297–311.

Wennerstrom, A. and Siegel, A.F. 2003. "Keeping the floor in multiparty conversations: Intonation, syntax, and pause". *Discourse Processes* 36(2): 77–107.

Wichmann, A. 2000. *Intonation in Text and Discourse*. London: Longman.

Authors' addresses

Salvatore Attardo
Department of Literature and Languages
Texas A&M University — Commerce
Commerce, TX 75429
USA

Salvatore_Attardo@tamu-commerce.edu

Lucy Pickering
Department of Literature and Languages
Texas A&M University — Commerce
Commerce, TX 75429
USA

Lucy_Pickering@tamu-commerce.edu

Amanda Baker
Department of Applied Linguistics & ESL
Georgia State University
P.O. Box 4099
Atlanta, GA 30302-4099
USA

eslambx@langate.gsu.edu

About the authors

Salvatore Attardo is Interim Dean of the College of Arts and Sciences at Texas A&M University — Commerce. He has written two books on humor and edited several collections of articles on the subject. He is the Editor-in-Chief of *HUMOR: International Journal of Humor Research*. He has also published numerous articles in semantics, pragmatics, and socio-linguistics.

Lucy Pickering is Associate Professor in the Department of Literature and Languages at Texas A & M-Commerce. She received her PhD from the University of Florida in 1999. Her research interests include intonation in discourse and the cross-linguistic transfer of prosodic features.

Amanda Baker is a PhD student in the Department of Applied Linguistics and ESL at Georgia State University. Her areas of interest include Second Language Listening, Speaking and Pronunciation, Teacher Education and Qualitative Research Methods.